Prof. Pietro Cerveri

AI METHODS for BIOENGINEERING CHALLENGES - 2025-26

Deep Learning for ECG Signal

Clinical context: In sports medicine, ensuring that athletes are fit to practice safely is a key aspect of cardiovascular prevention. During medical assessments for sport eligibility, **electrocardiography (ECG)** plays a central role in detecting early signs of cardiac abnormalities that may increase the risk of sudden cardiac events during physical activity. By recording the electrical activity of the heart through electrodes placed on the body surface, the ECG provides essential information on rhythm, conduction, and repolarization. Its analysis allows clinicians to identify conditions such as arrhythmias, conduction defects, or hypertrophic patterns that may contraindicate intense exercise. However, the interpretation of ECGs can be challenging due to signal variability and the subtle nature of some pathological patterns. The use of **deep learning techniques** represents a promising approach to automatically extract relevant features and improve the detection of at-risk individuals, supporting a more accurate and objective decision process for sport eligibility.

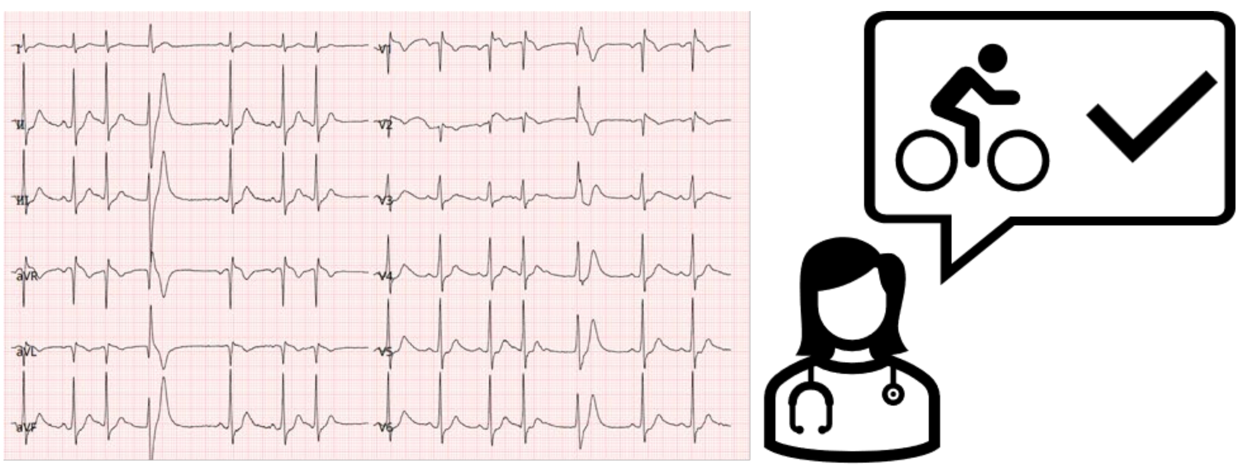

Figure 1. ECG picture with depicted traces.

Dataset

The dataset used in this project originates from the VALETUDO study conducted at IRCCS Ospedale Galeazzi Sant'Ambrogio. It comprises standard 12-lead electrocardiographic (ECG) recordings, stored in MATLAB format, from 526 subjects undergoing cardiovascular evaluation for sports eligibility, including exercise stress testing.

Each ECG recording has a total duration of 10 seconds (see traces in Figure 1) and is acquired at a sampling frequency of 1000 Hz. Due to the acquisition protocol of the recording system, the 12-lead ECG is provided as two consecutive sets of six synchronized leads: the first set includes the limb leads (I, II, III, aVR, aVL, aVF) and the second set includes the precordial leads (V1–V6), each recorded over a 5-second interval. As a consequence, a small temporal offset may exist between the two lead sets, which is accounted for during signal processing (see Figure 2 below).

In addition to the ECG signals, a tabular dataset in Microsoft Excel format is available and includes demographic, anthropometric, and clinical variables such as age, weight, height, sex, training load, and sport classification, together with a binary label indicating sports eligibility.

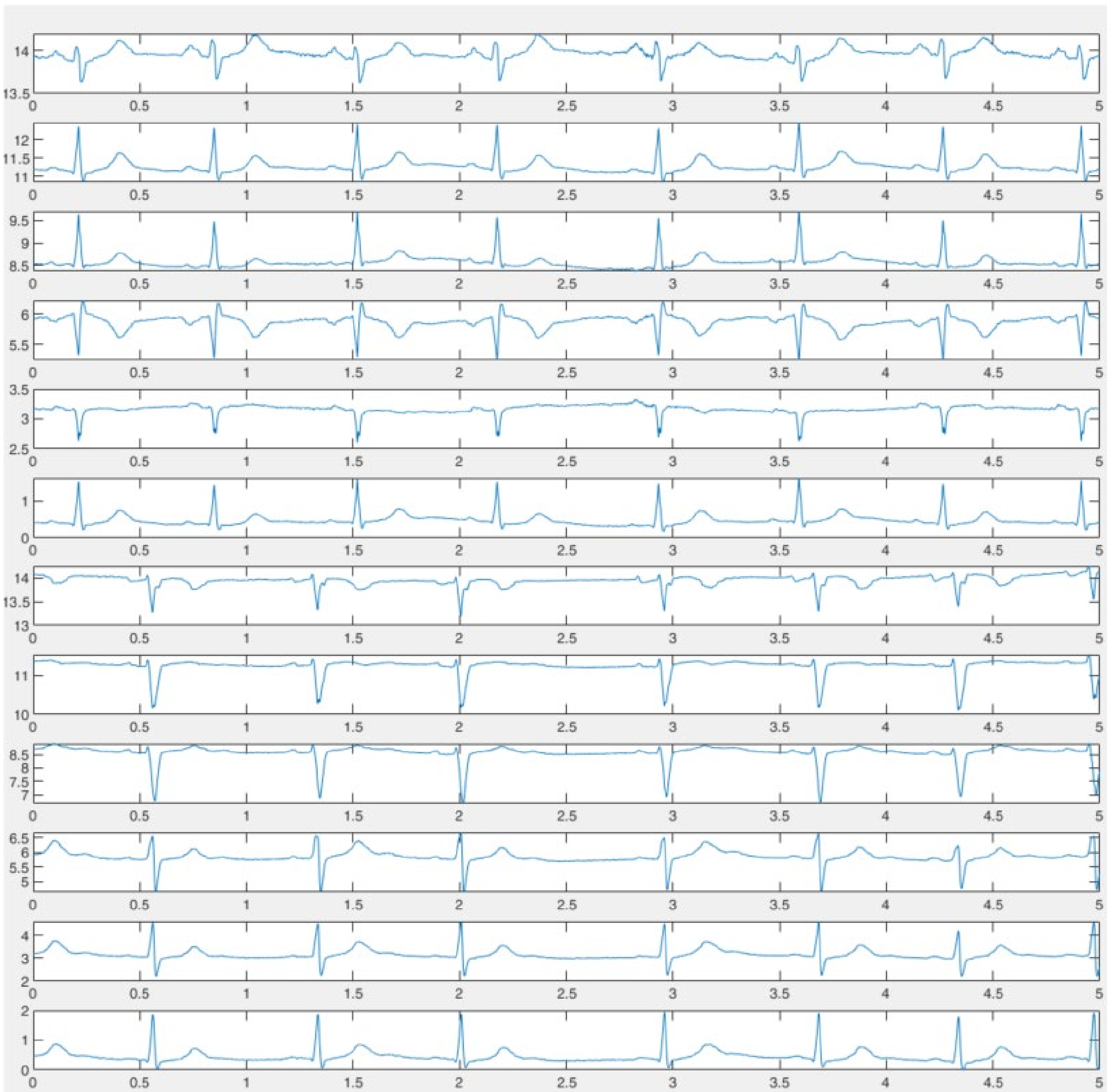

Figure 2. 12 Leads stacked. Notice the lag between the limb leads with respect two the

Aim

End-to-end deep learning methods to process ECG and perform classification according to sport ability label

Using deep neural networks, you implement models to classify patients according to their sport eligibility status. Different architectures may be tested and compared. The input will primarily consist of ECG signals, but tabular data may be integrated as additional features. Proper train/validation/test splitting and possibly cross-validation are essential to ensure robust evaluation.

Indicative processing steps

1. Data analysis (sampling frequency, lead proper ordering)
2. Pre-processing (baseline removal, filtering, normalization)
3. Training data setup (signal array arrangement, time period, signal stacking, label balancing)
4. Processing architecture (custom/literature-based network,
5. Learning strategy (optimizer, learning rate, batch size, early stopping)
6. Test metrics

Data analysis: Consider that each ECG is composed of two sets of 6 synchronized leads of 5 seconds each

Preprocessing and data cleaning: As in any real-world dataset, a first step will consist of cleaning and preprocessing. Verify the need of removing subjects with missing or inconsistent data (outliers, incomplete labels), imputing missing tabular values if needed, and applying signal preprocessing techniques such as filtering, baseline correction, and signal scaling.

Training data setup: You have to decide how to use to the two sets of leads. One possibility is to stack together the 12 leads so that you have 5000x12 as a input to the network. However, recall that there a lag in between the two sets. You may evaluate to implement the alignment by 1) estimating the temporal lag of the second wrt to the first set, using R peaks for instance, 2) shift signals and padding for missing samples. This stacking strategy preserves the full lead-specific morphological information while enabling the model to learn spatial and temporal patterns across multiple leads within a unified framework The second possibility is to consider two sets as independent. This would require a two-branch encoder for you deep network, ending up into a common bottle neck.

Data analysis: Each ECG recording is provided as two consecutive sets of six synchronized leads, each lasting 5 seconds (limb leads followed by precordial leads).

Preprocessing and data cleaning: As with any real-world dataset, the first step is data cleaning and preprocessing. This includes checking for missing or inconsistent entries (e.g., outliers, incomplete labels), removing subjects if necessary, and imputing missing tabular values when appropriate. For ECG signals, standard preprocessing should be applied, such as bandpass filtering, baseline-wander correction, and amplitude scaling/normalization to reduce inter-subject variability.

Training data setup: A key design choice is how to use the two 6-lead blocks.

> **(1) Stacked 12-lead input:** Each ECG consists of two consecutive 5-second segments sampled at 1000 Hz, each containing 6 synchronized leads (5000 samples/lead). One option is to stack the two segments into a 12-channel input tensor of shape 5000×12, where the limb leads and precordial leads are treated as separate channels over the same nominal time axis. Because the acquisition introduces a temporal lag between the two 6-lead blocks, alignment can be performed by (i) estimating the relative delay of the second block with respect to the first (e.g., using R-peak detection and matching or cross-correlation), and (ii) shifting the precordial leads accordingly with appropriate padding/truncation. This preserves lead-specific morphology while enabling the model to learn multilead spatiotemporal patterns.
>
> **(2) Two-branch input:**, the two 6-lead blocks can be treated as independent inputs, using a two-branch encoder that processes each block separately and fuses their representations into a shared bottleneck before classification, avoiding the assumption of perfect simultaneity across all 12 leads.

Processing architecture: The task is to design a deep learning architecture that classifies sports eligibility directly from ECG signals. The network should take preprocessed ECG data as input and learn discriminative features in an end-to-end manner. Students may choose between two main strategies:

(i) a single-branch architecture, where the two 6-lead ECG blocks are aligned (if necessary) and stacked into a 12-channel input processed by a 1D convolutional encoder; or

(ii) a multi-branch architecture, where each 6-lead block is processed independently by a dedicated encoder, and the resulting feature representations are fused in a shared bottleneck before the final classification layer.

In both cases, the architecture should be designed to capture temporal patterns within each lead and relationships across leads, while keeping the model complexity appropriate for the dataset size.

Learning strategy: The network is trained using a supervised learning approach. An adaptive optimizer such as Adam or AdamW is recommended, with an initial learning rate typically in the range of 1e-4 to 1e-3. The batch size should be chosen according to GPU memory constraints (commonly between 16 and 64 for ECG signals). To prevent overfitting, early stopping should be applied by monitoring the validation loss or a performance metric such as AUROC, with training stopped if no improvement is observed for a fixed number of epochs (e.g., 10–15). Learning rate scheduling (e.g., step decay or cosine annealing) may be used to improve convergence stability.

Test metrics: Model performance should be evaluated on a held-out test set using metrics appropriate for binary classification. In addition to overall accuracy, threshold-independent metrics such as the area under the ROC curve (AUROC) should be reported to assess the model's discriminative ability. Given the screening nature of sports eligibility, sensitivity (recall) and specificity are particularly important to quantify the ability to correctly identify non-eligible and eligible subjects, respectively. The F1-score can be used to balance precision and recall, especially in the presence of class imbalance. Confusion matrices should be reported to provide a clear interpretation of classification errors.

Use of anamnestic patient data: In addition to ECG signals, selected anamnestic and clinical variables (e.g., age, sex, anthropometric measures, training load, and sport classification) may be incorporated to enhance model performance. For instance, these variables can be processed through a dedicated fully connected branch and fused with ECG-derived features at a later stage of the network.